

DATA TRANSFER DEVICE BY MULTI-BUFFER CONTROL IN PARALLEL COMPUTERS

Publication number: JP10307802

Publication date: 1998-11-17

Inventor: MAYA YUZURU; OTSUJI AKIRA; SHIGENO TAKEYUKI; SAKURABA TATSUTOSHI

Applicant: HITACHI LTD

Classification:

- international: G06F13/14; G06F15/163; G06F13/14; G06F15/16; (IPC1-7): G06F15/163; G06F13/14

- european:

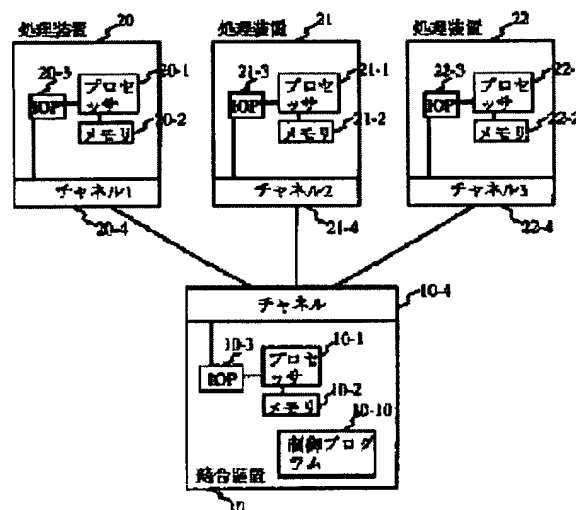
Application number: JP19970119089 19970509

Priority number(s): JP19970119089 19970509

Report a data error here

Abstract of JP10307802

PROBLEM TO BE SOLVED: To shorten command response time by transferring data after a command processing is ended in the write command processing of a coupling device. **SOLUTION:** A system is constituted of the coupling device 10 and three processors 20-22 and the hardware constitution of the coupling device 10 and the processors 20-22 is the same. All the processors 20-22 issue write and read commands to the coupling device 10 and access the coupling device 10. The coupling device 10 performs the priority control of the processing of a command main body and the processing of data transfer so as to free at least one command area and a data area for all channels 20-4-22-4 so as to receive commands from all the channels 20-4-22-4 at all times. That is, when the coupling device 10 writes data from the processors 20-22 by the command processing, the coupling device 10 returns a response to the processors 20-22 and then, transfers them from a transfer data storage area to memories 20-2-22-2.



Data supplied from the esp@cenet database - Worldwide

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平10-307802

(43)公開日 平成10年(1998)11月17日

(51)Int.Cl. ⁸	識別記号	F I
G 0 6 F 15/163		G 0 6 F 15/16 3 1 0 V
13/14	3 1 0	13/14 3 1 0 Y

審査請求 未請求 請求項の数6 O L (全 5 頁)

(21)出願番号 特願平9-119089

(22)出願日 平成9年(1997)5月9日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 真矢 誠

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(72)発明者 大辻 彰

神奈川県横浜市戸塚区戸塚町5030番地 株

式会社日立製作所ソフトウェア開発本部内

(72)発明者 茂野 丈至

神奈川県秦野市堀山下1番地 株式会社日

立製作所汎用コンピュータ事業部内

(74)代理人 弁理士 小川 勝男

最終頁に続く

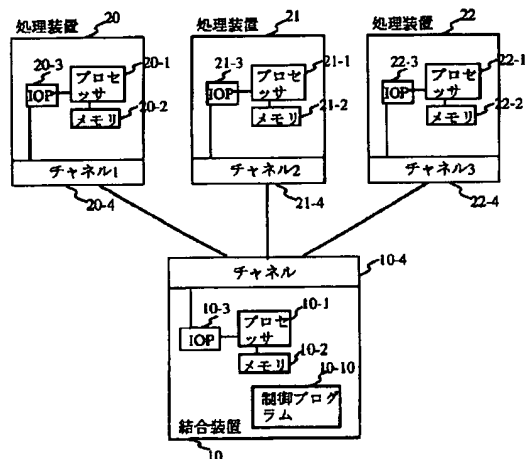
(54)【発明の名称】 並列計算機におけるマルチバッファ制御によるデータ転送装置

(57)【要約】

【課題】本発明の課題は、処理装置と結合装置から構成される並列計算機において、結合装置の書き込みコマンド処理のデータ転送時間を削除し、処理装置からコマンドを受信してから応答を戻すまでのコマンド応答時間を短縮することである。

【解決手段】本発明では、結合装置は処理装置から受信するデータを格納するためのバッファを複数個用意する。結合装置は、処理装置から受信したコマンドが書き込みコマンドの場合、データを管理するデータ項目にデータのアドレスを示す領域を設け、ここにデータアドレスを設定し、コマンドの応答を処理装置に戻す。その後、結合装置はデータをバッファからメモリに転送する。このように、本発明における結合装置の書き込みコマンド処理では、コマンド処理終了後にデータ転送を行うため、コマンド応答時間は短縮できる。

図 1



【特許請求の範囲】

【請求項1】複数の処理装置とこれらが共有する結合装置から構成される並列計算機において、結合装置と処理装置は、チャンネル対応にコマンドとデータを格納する複数のメッセージバッファを格納する領域を設けることを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【請求項2】請求項1において、結合装置はコマンド処理終了後に、メッセージバッファからメモリにデータ転送を行うことを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【請求項3】請求項1において、結合装置は全てのチャンネルにおけるメッセージバッファの空き数が少なくとも1以上であるようにデータ転送を行うことを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【請求項4】請求項1において、結合装置は、処理装置からデータ転送完了前にデータの読み出しコマンドを受信すると、メッセージバッファから読み出すことを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【請求項5】請求項1において、チャンネル毎に1個のコマンド領域とデータ領域を用意し、残りはチャンネル共通のメッセージバッファを設け、全てのチャンネルで共有することを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【請求項6】請求項1において、全てのメッセージバッファが使用中ならば、コマンドは受信できないことを特徴とする並列計算機におけるマルチバッファ制御によるデータ転送装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、処理装置とこれらが共有する結合装置から構成される並列計算機に係わり、結合装置のコマンド処理におけるデータ転送時間を削除し、処理装置のコマンド応答時間を短縮させる方法に関する。

【0002】

【従来の技術】処理装置とこれらが共有する結合装置から構成される並列計算機は、特開平6-4490「データ処理システム」に記載されている。これによると、処理装置（中央処理複合体：CPC）はメッセージバッファにデータを格納して、結合装置（SES機構）にデータを転送している。

【0003】リスト構造の読み出しと書き込みは、特開平6-83783「条件付きオペレーション提供装置及び方法」に記載されているように、リスト入力を読み出されている時、リスト出力が書き込まれている時、コマンド処理中にデータを転送している。

【0004】このため、従来の方式では、データサイズ

が大きくなると、コマンド処理におけるデータ転送時間が長くなり、コマンドの実行時間が長くなるという問題がある。

【0005】

【発明が解決しようとする課題】処理装置は結合装置にデータを書き込む場合、処理装置は結合装置に書き込みコマンドとデータを送信する。

【0006】結合装置は書き込みコマンドを受信すると、このコマンドを解析し、データ項目を作成する。そして、メッセージバッファから結合装置のメモリにデータを転送する。最後に、結合装置は応答を処理装置に戻す。

【0007】この結果、結合装置はコマンド処理中にデータ転送を行うため、コマンドの実行時間が長くなるという問題がある。コマンド応答時間は、コマンド本体の処理時間とデータ転送時間からなる。本発明の課題は、結合装置のコマンド処理において、データ転送時間を削除し、処理装置からコマンドを受信してから応答を戻すまでのコマンド応答時間を短縮することである。

【0008】

【課題を解決するための手段】本発明では、処理装置と結合装置から構成される並列計算機において、結合装置は処理装置から受信するデータを格納するためのバッファを複数個用意する。

【0009】結合装置は、処理装置から受信したコマンドを解析する。これを書き込みコマンドの場合、結合装置はデータを管理するデータ項目にデータのアドレスを示す領域を設け、ここにデータアドレスを設定し、コマンドの応答を処理装置に戻す。その後、結合装置はデータをバッファからメモリに転送する。

【0010】このように、本発明における結合装置の書き込みコマンド処理では、コマンド処理終了後にデータ転送を行うため、処理装置からコマンドを受信してから応答を戻すまでのコマンド応答時間は短縮できる。

【0011】

【発明の実施の形態】以下、本発明の実施の形態を示す。

【0012】図1は、本発明による並列計算機のシステム構成図である。本発明によるシステム構成は、結合装置（10）と3台の処理装置（20～22）から構成する。結合装置（10）と処理装置（20～22）のハードウェア構成は同一とし、それぞれプロセッサ（10-1～22-1）、メモリ（10-2～22-2）、IOP（入出力プロセッサ：Input Output Processor）（10-3～22-3）、およびチャンネル（10-4～22-4）から構成する。処理装置（20）はチャンネル1（20-4）を、処理装置（21）はチャンネル2（21-4）を、処理装置（22）はチャンネル2（22-4）使用する。

【0013】そして、全ての処理装置（20～22）は

結合装置(10)に書き込みコマンドおよび読み出しを発行することにより、結合装置(10)をアクセスする。図2は結合装置のデータ管理方法を示す図である。データを書き込む場合、データ項目(31、32)を作成し、待ち行列で管理する。待ち行列毎にデータ項目制御(30)を設ける。

【0014】データ項目制御(30)には、最初に接続するデータ項目(31)のアドレスを格納する。

【0015】データ項目(31)は、データ識別子(31-1)、データポインタ(31-2)、データ位置フラグ(31-3)および次データ項目ポインタ(31-4)から構成される。データ項目(32)の内容はデータ項目(31)と同一である。

【0016】データ識別子(31-1)は処理装置(20~22)がデータを識別するためのものであり、ユニークな値である。

【0017】データポインタ(31-2)はデータのアドレスを示す。

【0018】データ位置フラグ(31-3)は、データが転送データ格納領域にあるか、あるいはメモリにあるかを示すフラグである。以下、これが"1"の場合には、データが転送データ格納領域にあることを、"0"の場合には、データが転送データ格納領域(50)からメモリ(20-2、21-2、22-2)に転送されたことを示す。

【0019】次データ項目ポインタ(31-4)は、次のデータ項目(32)のアドレスを格納する。

【0020】図3は結合装置の転送データ格納領域を示す図である。

【0021】結合装置(10)の転送データ格納領域(50)はチャンネル1(20-4)に対して、 n 個のコマンドとデータが格納できるように、コマンドの領域(50-1~5n-1)、データの領域(50-2~5n-2)および応答の領域(50-0)を設ける。

【0022】チャンネル2(21-4)とチャンネル3(22-4)も、同様にコマンドの領域(60-1~6n-1、70-1~7n-1)、データの領域(60-2~6n-2、70-2~7n-2)および応答の領域(50-0)を設ける。

【0023】これにより、結合装置(10)は転送データ格納領域(50、60、70)からメモリ(20-2、21-2、22-2)にデータ転送している間でも、空き領域があれば、結合装置(10)は処理装置(20~22)からコマンドを受信できるようになる。

【0024】図4は結合装置の書き込み処理手順を示す図である。処理装置(20)は書き込みコマンドを送信するために書き込みコマンドを発行し(処理100)、このコマンドとデータを送信する(処理101)。

【0025】一方、結合装置(10)はこのコマンドを受信すると、転送データ格納領域(50)からコマンド

(50-1)とデータ(50-2)を読み出し、コマンドを解析する(処理102)。そして、受信したデータを管理するためのデータ項目(31)を作成する(処理103)。そして、作成したデータ項目(31)を以下のように、設定する(処理104)。

【0026】まず、結合装置(10)はユニークなデータ識別子(31-1)を設定し、データ位置フラグ(31-3)は、まだデータがメモリ(20-2)に転送されていないため"1"を設定し、データポインタ(31-2)は転送データ格納領域(50)に格納されているデータのアドレスを設定し、最後にこのデータ項目(31)を指定された場所に接続するために次データ項目ポインタ(31-4)を設定する。

【0027】更に、結合装置(10)は応答を作成し(処理106)、転送データ格納領域(50)の応答(50-0)に格納する。

【0028】最後に、結合装置(10)は処理装置(20)に応答(50-0)を送信する(処理107)。

【0029】その後、結合装置(10)はデータ(50-2)を転送データ格納領域(50-0)からメモリ(20-2)に転送し、このデータ転送が完了すると、データ項目(31)のデータポインタ(31-2)を再設定し、データ位置フラグ(31-3)は"1"から"0"に更新する。図5は結合装置の読み出し処理手順を示す図である。処理装置(21)は読み出しコマンドを送信するために読み出しコマンドを発行し(処理200)、このコマンドを送信する(処理201)。

【0030】一方、結合装置(10)はこのコマンドを受信すると、転送データ格納領域(60)からコマンド(60-1)を読み出し、コマンド(60-1)を解析する(処理202)。そして、データ識別子(31-1)から該当するデータ項目(31)を検索する(処理203)。そして、検索したデータ項目(31)のデータポインタ(31-2)からデータを読み出し、転送データ格納領域(60)にデータ(60-2)を設定する(処理204)。

【0031】更に、転送データ格納領域(60)に応答(60-0)を作成する(処理205)。転送データ格納領域の応答に格納する。最後に、結合装置(10)は処理装置(21)に応答(60-0)とデータ(60-2)を送信する(処理206)。

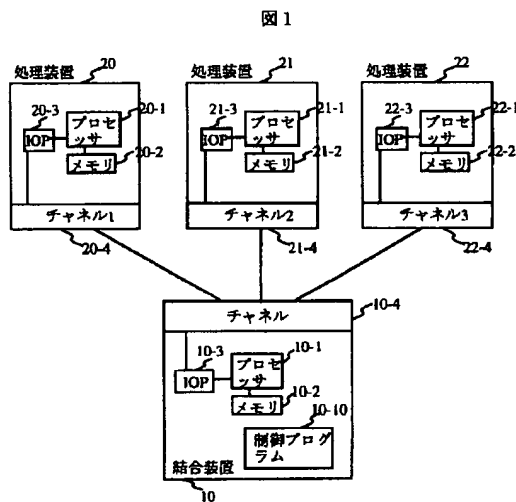
【0032】上記の実施例におけるコマンド処理は、コマンド本体の処理とデータ転送の処理に分けられ、コマンド本体の処理を優先させる。このため、全てのコマンド領域(50-1~5n-1、60-1~6n-1、70-1~7n-1)とデータ領域(50-2~5n-2、60-2~6n-2、70-2~7n-2)を使用するケースがある。このケースには、結合装置(10)は処理装置(20~22)からのコマンドが受信できなくなる。このようなケースをなくすため、結合装置(1

0)は全てのチャンネル(20-4~22-4)からコマンドがいつも受信できるようにするため、全てのチャンネル(20-4~22-4)について、少なくとも1個のコマンド領域とデータ領域を空けておくように、コマンド本体の処理とデータ転送の処理の優先制御を行う。

【0033】上記の実施例では、チャンネル(20-4~22-4)毎にn個のコマンド領域(50-1~5n-1、60-1~6n-1、70-1~7n-1)とデータ領域(50-2~5n-2、60-2~6n-2、70-2~7n-2)を設ける方式で説明した。この他に、チャンネル(20-4~22-4)毎に1個のコマンド領域とデータ領域を用意し、残りはチャンネル(20-4~22-4)共通のコマンド領域とデータ領域を設け、全てのチャンネル(20-4~22-4)で共有することも可能である。

【0034】

【図1】



*【発明の効果】本発明では、複数の処理装置とこれらが共有する結合装置から構成される並列計算機において、結合装置はコマンド処理により処理装置からのデータを書き込む際、結合装置は処理装置に応答を返した後に、転送データ格納領域からメモリに転送することにより、コマンド応答時間を短縮する。

【図面の簡単な説明】

【図1】本発明による並列計算機のシステム構成図である。

10 【図2】結合装置のデータ管理方法を示す図である。

【図3】結合装置の転送データ格納領域を示す図である。

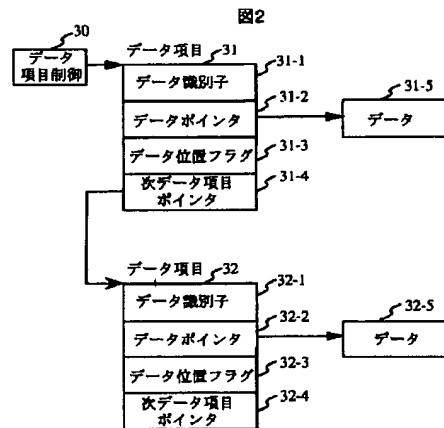
【図4】結合装置の書き込み処理手順を示す図である。

【図5】結合装置の読み出し処理手順を示す図である。

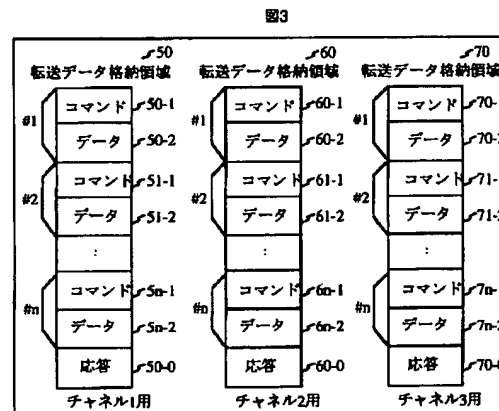
【符号の説明】

* 10…結合装置、20~22…処理装置

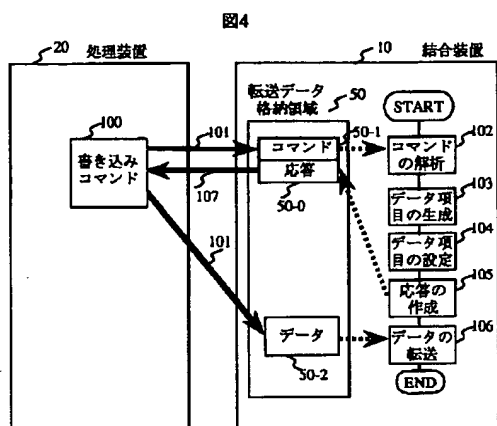
【図2】



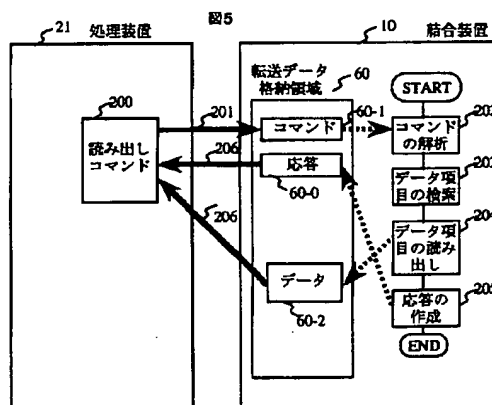
【図3】



【図4】



【図5】



フロントページの続き

(72)発明者 櫻庭 健年

神奈川県川崎市麻生区王禅寺1099番地 株
式会社日立製作所システム開発研究所内